

Figure 1: Re-Using Columns that Correspond to a Shared Prefix. Note that, for the columns that correspond to a shared prefix, the values are the same in both grids. (This example uses the simple edit distance detailed in the specification text.)

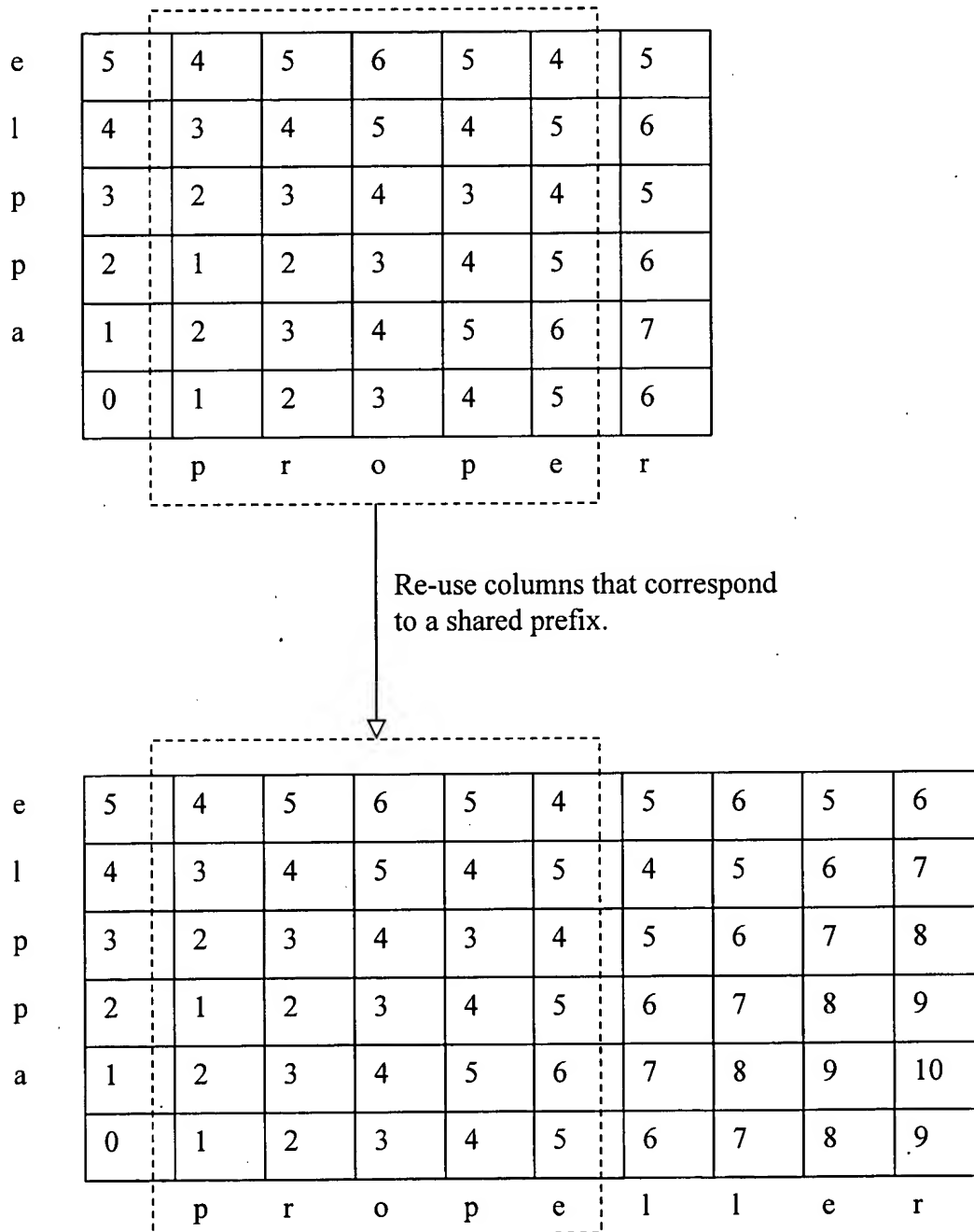


Figure 2: Preprocessed Text List. The index of the first different character from the previous text supports column reuse for shared prefixes. The list of (prefix, index of next text lacking the prefix) supports avoiding computation for texts that share a prefix that causes edit distance to exceed a threshold. (A prefix occurs in a list if it is not shared with the previous text and is shared with the subsequent text.)

Text List Index	Text	First Different Character	List of (Prefix, Index of Next Text Lacking Prefix)
0	arm	0	(a, 2), (ar, 2), (arm, 2)
1	arms	3	
2	banana	0	(b, 5), (ba, 5), (ban, 4)
3	band	3	
4	bar	2	
5	cab	0	(c, 7)
6	club	1	

Figure 3: Computing Column Bands. The diagram illustrates which grid values are computed for a threshold value of four. Using the strategy of computing column bands, only the cells shown with values are computed. For each column except the last, the circled cells are the range of cells from the bottom cell with value less than the threshold to the top cell with value less than the threshold. Cells to the right of cells in this range are computed normally. So is the cell above the cells to the right. Then cells above are computed until a value is at least the threshold.

s					4	5
e				4	3	4
l			4	3	2	3
p	4	5	4	3	2	3
m	3	4	3	2	3	4
i	2	3	2	3	4	
d	1	2	3	4	5	
	0	1	2	3	4	
	s	i	m	p	l	y

Figure 4: Re-Using Columns that Correspond to Recurrence-Preserving Characters. In case A, the first two columns correspond to a shared prefix. In case B, in each text the column corresponds to a character not found in the search string, i.e., neither “s” nor “p” occurs in “brink.” In case C, the column corresponds to the same character in each text. In case D, in each text the column corresponds to a character not found in the search string. In each case, the columns from the top computation are the same as in the bottom computation, so they can be re-used without being re-computed.

